How to tell what is going on (or not going on)

not search but keeping track

> Jussi Karlgren
it's not search any more
it's not search any more

it's not about a needle in a haystack
it's not search any more

it's not about a needle in a haystack

it's about the shape of the wheatfield over time
it's not search any more

it's not about a needle in a haystack

it's about the shape of the wheatfield over time

it's about information awareness, not misplaced items
big data is not just slightly larger data
it's not a scalability issue
it's about information awareness, not search
the place of text analytics
the slightly more complex picture
the technology enabler: the big data stack
a semantic base technology?

> is this an example of that?
> are these two the same?
> has this changed? how?
> what is the relation of this and that?
> is this a new way of saying that?
> are these or those more like this?
> is this typical or strange?
> can we trust this?
> does the author believe this to be true?
we build a living lexicon
we build a living lexicon
we build a living lexicon
we build a living lexicon
we build a living lexicon
three points i want to address here, in terms of application

> language coverage

> "sentiment" analysis

> making sense of large amounts of text
evaluating marketing
evaluating marketing

traditional marketing campaign
evaluating marketing

traditional marketing campaign

giveaways to cosmetics subscribers
evaluating marketing

giveaways to cosmetics subscribers
giveaways to bloggers

campaign

traditional marketing

evaluating marketing

what was said is not what is interesting. Volume and tonality is.

giveaways to cosmetics subscribers
giveaways to bloggers

traditional marketing campaign
tracking violence in the world
tracking violence in the world

Kazakhstan stands out - unreported in Western media, a violent altercation in a court where protestors were sentenced to prison sentences took place on this week in 2012.
unrest or distrust?
the internet is not in english any more

the internet is not in english any more

the internet is not in english any more

the internet is not in english any more
swedish
finnish
danish
norwegian
english
french
german
portuguese
russian
...
...
how great is the cost of adding new languages?
how great is the cost of adding new languages?

should the channels be separated? (given code switching?)
sentiment analysis?
And the sound quality - my God!
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been an expansion pack.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been a expansion pack.

I loved The Spy Who Came In From The Cold but the movie is a bit dated in a way the book never will be.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been a expansion pack.

I loved The Spy Who Came In From The Cold but the movie is a bit dated in a way the book never will be.

Meat is more environmentally friendly than seafood.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been a expansion pack.

I loved The Spy Who Came In From The Cold but the movie is a bit dated in a way the book never will be.

Meat is more environmentally friendly than seafood.

I am unsure about the feasibility of this knitting pattern.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been a expansion pack.

I loved The Spy Who Came In From The Cold but the movie is a bit dated in a way the book never will be.

Meat is more environmentally friendly than seafood.

I am unsure about the feasibility of this knitting pattern.

I love the Samsung B2710 but I would not recommend it to my colleagues.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been a expansion pack.

I loved The Spy Who Came In From The Cold but the movie is a bit dated in a way the book never will be.

Meat is more environmentally friendly than seafood.

I am unsure about the feasibility of this knitting pattern.

I love the Samsung B2710 but I would not recommend it to my colleagues.

I don't know if I should call her up – I liked her when I met her last weekend.
And the sound quality - my God!

Raymond left no room for error on his recordings and it shows.

Definitely one of the better tracks on the album.

Wow, could have been a expansion pack.

I loved The Spy Who Came In From The Cold but the movie is a bit dated in a way the book never will be.

Meat is more environmentally friendly than seafood.

I am unsure about the feasibility of this knitting pattern.

I love the Samsung B2710 but I would not recommend it to my colleagues.

I don't know if I should call her up – I liked her when I met her last weekend.

This is true.
joy
joy  jealousy

joy
joy

fear

skepticism

jealousy
hate
fear
desire
anger
skepticism
love
joy
excitement
surprise
jealousy
happiness
love
desire
fear
hate
skepticism
anger
joy
excitement
surprise
jealousy
happiness
excitement

joy

love

desire

fear

hate

skepticism

anger

surprise

happiness

jealousy
example: gamer gate tweets

Possibly the most offensive misrepresentation I've seen

So @ChrisWarcraft is speaking my law class on Thursday. I was excited, but after #GamerGate, I am now holymother-of-all-creation excited

@ChrisWarcraft: Just finished up a piece on #Gamergate that I'm super excited about! Got to think up a whole bunch of new insults.

Oh man I love to done spiked armor with nipples and fight for love and honor. Thanks for the compliment #GamerGate

...
Possibly the most offensive misrepresentation I've seen

So @ChrisWarcraft is speaking my law class on Thursday. I was excited, but after #GamerGate, I am now holy-mother-of-all-creation excited

@ChrisWarcraft: Just finished up a piece on #Gamergate to think up a whole

Oh man I love to done spiked armor with nipples and fight for love and honor. Thanks for the compliment #GamerGate

...

sentiment analysis? 10%
coverage wrt to POS-NEG
example: gamer gate tweets

Possibly the most offensive misrepresentation I've seen

So @ChrisWarcraft is speaking my law class on Thursday. I was excited, but after #GamerGate, I am now holy-mother-of-all-creation excited!

@ChrisWarcraft: Just finished up a piece on #Gamergate and I got to think up a whole bunch of new insults.

Oh man I love to done spiked armor with nipples and fight for love and honor. Thanks for the compliment!

sentiment analysis? 10% coverage wrt to POS-NEG

multi-dimensional tonality: > 50% coverage
utterance clustering ≈ stories API
EBOLA

Situation de la lutte contre la maladie à virus Ebola a...

Rom (dpa) - Ein italienischer Arzt hat sich in Sierra L...

Ett SAS-plan från Paris isolerades på den danska flygpl...

RT @ZahiraXO: ZOMBIE APOCALYPSE : Ebola victims in...
Mali confirms two new Ebola cases, India News, India News Headlines, India News Updates, Latest News about India, Breaking News

- "Mali confirms two new Ebola cases - Business Standard" http://t.co/AnLaj21FvY

Ebola isolation at US base 'pretty much vacation'

- Italy's first Ebola victim is flown into Rome in specially sealed plane
- troops have been deployed to Liberia and Senegal for the Ebola mission, along with almost 100 contractors and nearly 60 Defense Department civilians.

Sierra Leone Ebola workers dump bodies in protest

- The head of the district Ebola Response Team, Abdul Wahab Wan, said the bodies had included those of two babies, and that some had been displayed around the hospital.
- "Displaying corpses in a very, very inhumane manner is completely unacceptable," said the spokesman for the National Ebola Response Centre, Sidi Yahya Tunis.
three points, again:

> language coverage - simple (ought to be!), but how to use?

> "sentiment" analysis - more complex than assumed previously!

> making sense of large amounts of text - largely presentation issue, especially given quality of docs!